Microchemical Journal 181 (2022) 107834

Contents lists available at ScienceDirect

Microchemical Journal

journal homepage: www.elsevier.com/locate/microc

Determination of total phenolic content and selected phenolic compounds in sweet wines by fluorescence spectroscopy and multivariate calibration

Michaela Jakubíková *, Jana Sádecká, Katarína Hroboňová

Institute of Analytical Chemistry, Faculty of Chemical and Food Technology, Slovak University of Technology, Radlinského 9, 812 37 Bratislava, Slovak Republic

ARTICLE INFO

Keywords:
Fluorescence spectroscopy
HPLC
Chemometrics
Wine
Total phenolic content
Phenolics

ABSTRACT

The aim of this work was to develop a simple, fast and reagent-free method for the prediction of total phenolic content (TPC) and selected phenolic compounds in Slovak botrytized Tokaj wines based on spectral data and partial least squares (PLS) regression. The significant variables for the prediction using PLS were obtained by evaluating the variable importance in projection (VIP). Folin-Ciocalteu and HPLC, respectively, were used as reference methods for the determination of TPC and selected phenolic compounds. Using UV–vis and excitation-emission matrix (EEM) fluorescence spectra recorded on the bulk and diluted samples, the calibration models for TPC were developed and compared. The best PLS model with relative predictive deviation (RPD) of 5.8, coefficient of determination (R2) of 0.972 and root mean squares error (RMSE) of 20.2 mg /L was based on the variables selected from unfolded EEM fluorescence spectra of diluted samples. Using unfolded EEM fluorescence spectra recorded on the diluted samples and VIP, the PLS calibration models for gallic, protocatechic, caffeic and *p*-coumaric acids and (+)-catechin were obtained with RPD > 4.0 and R2 > 0.9. The RMSE values were 0.7, 0.5, 0.3, 0.2 and 1.0 mg/L for gallic, protocatechic, caffeic and *p*-coumaric acids and (+)-catechin, respectively.

1. Introduction

Wine, an alcoholic beverage obtained from fermented grape juice, has been produced for millennia, resulting in the diverse styles of wine currently available. Wine styles are often related to the unique geographical and climatic conditions in which they originated. Botrytized wines are rarely produced in the regions where geographical and climatic conditions allow the fungus *Botrytis cinerea* to produce a specific fungal infection of the grape called noble rot [1,2]. In the Tokaj region, by mixing noble rotten grape berries and grape fermenting must or young wine (with a sugar content of at least 21 kg/100 L), a high-quality botrytized wine called Tokaj selection is produced. The quality and price of the Tokaj selection depends on the number of butts added to 136 L of must, where 1 butt represents approximately 25 kg of noble rotten grape berries. The wines are graded from 2 to 6 butts. Tokaj selection shall mature at least three years in a specific microclimate, which, together with the effects of the *Botrytis cinerea*, gives a typical aroma, taste and color to wine [3,4]. In addition, Tokaj selections must be made from strictly defined white wine grape varieties, Lindenblaetrige, Yellow Muscat and Furmint [3–5].

A total phenolic content (TPC) of white wine expressed as gallic acid (GAE) varies in the range of 180 to 280 mg GAE/L depending on the aging time [1]. Bajčan et al. [6] found that the TPC of Slovak Tokaj wines of varieties Lindenblaetrige, Yellow Muscat and Furmint was 490, 526 and 405 mg GAE/L, respectively. Compared to other white monovarietal wines, Tokaj wines had both higher TPC and higher antiradical activity. Because the primary source of phenolic compounds in most wines is berry skin, botrytized wines made by macerating berries in fermented must or young wine contain more or less increased amounts of phenolics depending on the duration of maceration compared to normal wines [1]. For example, Hungarian Tokaj botrytized wines contained twice the total phenolics compared to non-botrytized ones [7]. Ballová et al. [8] found the TPC higher in Slovak Tokaj 6 and 5 butt wines than in the non-botrytized wines. In addition to the high phenolic content, all Slovak Tokaj wines showed high antiradical activity (against DPPH radical) ranging from 58 % in Furmint to 78 % in 6 butt wine [9].

The most widely used and cited method for determining the TPC in food and beverages is spectrophotometric Folin-Ciocalteu (FC) method [10,11]. It is the official method of analysis of the Association of Official Analytical Chemists [12] for estimating of total soluble phenolic content in dietary supplements and the standard procedure adopted by the International Organisation of Vine and Wine (OIV) for wine analysis [13].

* Corresponding author.
E-mail address: michaela.jakubikova@stuba.sk (M. Jakubíková).

https://doi.org/10.1016/j.microc.2022.107834
Received 10 May 2022; Received in revised form 26 July 2022; Accepted 27 July 2022
Available online 30 July 2022

Although the method is widely accepted, it requires chemicals and time to complete the reaction, and solutions containing the FC reagent are toxic and must therefore be disposed of as hazardous waste. To overcome these drawbacks, the combinations of spectrometric methods with chemometrics to predict different physicochemical parameters, including TPC, are increasingly being investigated.

UV–vis spectroscopy was used in estimating the TPC in red wines [14,15], sparkling wine press juice fractions [16], white and red grape vinegers [17] and apple juices [18]. Compared to infrared (IR) spectroscopy, UV–vis spectroscopy provided better results (higher coefficient of determination (R2) and lower root mean squares error (RMSE)) for apple juices [18]. In contrast, the results obtained by both methods were similar for white and red grape vinegars, and fusion of data from both methods was preferred [17]. Because many of the components contributing to the TPC show fluorescence, another popular method is fluorescence spectroscopy, which provides a wide variety of input data to chemometric models: individual emission spectra [19], unfolded excitation-emission matrix (EEM) fluorescence spectra [20–22], selected pairs of excitation and emission wavelengths [20,21], individual synchronous fluorescence spectra [23] and fused synchronous fluorescence spectra [24]. Thus, fluorescence spectroscopy allowed the prediction of TPC in beer [24], cachaca [20], red wine [22), plum extract [19] and tea [21,23].

Unlike a variety of spectroscopic techniques, one chemometric method, partial least squares (PLS) regression, is preferred to develop calibration models between spectra and the TPC, or each of the other chemical parameters [14–24]. Five vinegar processing parameters including Brix, TPC, total flavonoid content (TFC), titratable acidity (TA) and pH were successfully predicted from PLS models of combined UV–vis and MIR datasets (R2 > 0.97) [17]. In a comparative study, PLS models with a good predictive ability were obtained for several apple juices quality parameters mostly by using a fusion of various spectra: soluble solids content (SSC) (UV–vis-NIR), TA (UV–vis-NIR), SSC/TA (Vis-NIR), pH (Vis-NIR), TPC (UV–vis) and TFC (UV) (R2 > 0.80) [18]. Considering emission spectra recorded at an excitation wavelength of 280 nm on plum extracts, PLS model provided good results about quantification of epicatechin with R2 value of 0.89 [19]. Synchronous fluorescence spectra obtained at wavelength interval of 20 nm associated with PLS were a rapid way to monitor TPC in non-fermented tea ((R2 > 0.96) [23]. Combining excitation-emission matrices with an unfolded PLS methodology, the selected models for TPC achieved R2 above 0.99 and 0.95 in aged cachaça [20] and tea [21] samples, respectively. The use of variable selection tool significantly decreased the number of variables and improved RMSE [20,21].

To summarize, information on the phenolic composition of botrytized wines is relatively rare and information on Slovak Tokaj wines is almost non-existent in the literature. The aim of this work was to develop a simple, fast and reagent-free method for the prediction of TPC as well as selected phenolic compounds in Slovak botrytized Tokaj wines using spectral data and PLS. Using UV–vis and unfolded EEM fluorescence spectra recorded on the bulk and diluted samples, the calibration models for TPC were developed and compared. Unfolded EEM fluorescence spectra of diluted samples were the basis for the determination of gallic, protocatechic, caffeic and *p*-coumaric acids and (+)-catechin. FC and HPLC, respectively, were used as reference methods for the determination of TPC and individual phenolic components.

2. Material and methods

2.1. Wine samples, standards and reagents

Forty-six botrytized wine samples from the Slovak Tokaj region were analyzed: 3 of two butt, 8 of three butt, 8 of four butt, 13 of five butt and 14 of six butt wines. Wines of the vintages ranging from 1959 to 2016 were obtained from three local producers (Tokaj wineries Ostrožovič spol. s r.o., Anna Nagyová-ZLATÝ STRAPEC and Tokaj & Co., s.r.o.). Samples were stored at 4 °C in dark and equilibrated at 20 °C before analysis. Bulk samples were analyzed by UV–vis and fluorescence spectroscopy and HPLC, while a 100-fold and 500-fold dilution was used to record UV–vis and fluorescence spectra, respectively. Water purified by a Milli-Q system (Millipore, USA) was used for all dilutions.

Caffeic acid, caftaric acid, gallic acid, *p*-coumaric acid, protocatechic acid, (+)-catechin, sodium carbonate and Folin–Ciocalteu reagent were purchased from Sigma Aldrich Chemie (Steinheim, Germany). Methanol and ethanol (HPLC gradient grade), acetic acid (99 %) were purchased from Merck (Darmstadt, Germany).

2.2. Determination of TPC by Folin-Ciocalteu method

The determination of TPC was based on a microscale protocol described by Waterhouse [11], using the Folin–Ciocalteu reagent [10] and gallic acid as standard. To prepare a stock solution of gallic acid, 0.5000 g of standard was dissolved in 10 mL of ethanol and then diluted to 100.0 mL with water. The stock solution was stored for one week under refrigeration at 4 °C. Gallic acid calibration solutions in the concentration range of 50 to 500 mg/L were daily prepared by diluting 1.0 mL to 10.0 mL of the stock solution into 100.0 mL with water. For the analysis, 20 μL of gallic acid calibration solution (50–500 mg/L) or wine solution (usually fourfold diluted bulk wine), 1.58 mL of water and 100 μL of Folin-Ciocalteu reagent were mixed and incubated 5 min. Then 300 μL sodium carbonate solution (20 %, w/v) was added and mixed, and after 2 hr incubation at room temperature, the absorbance of the mixture was measured at 760 nm with 10 mm glass cell (2-mL). Each calibration or wine solution was analyzed in triplicate and the total phenolic content was expressed as milligrams of gallic acid equivalents/liter of bulk wine (mg GAE/L wine). Bulk botrytized wine must be diluted with water (usually fourfold) to fall into the calibration range of the standard.

2.3. UV–vis and fluorescence spectroscopy

UV–vis absorption spectra of bulk (200–700 nm, each 1 nm) and diluted (200–500 nm, each 1 nm) wine samples were recorded using UV 1800 Spectrophotometer (Shimadzu, Japan), quartz cell of 10 mm, scanning speed of 200 nm/min and software UV PROBE 2.33. The three UV–vis spectra for each sample were averaged, and the mean spectra were used to chemometric analysis.

Fluorescence spectra were obtained using the Perkin-Elmer LS 50 Luminescence Spectrometer (Perkin-Elmer, USA), quartz cell of 10 mm, the widths of both the excitation slit and emission slits of 5 nm, a scanning speed of 200 nm/min and FL Data Manager Software for spectral acquisition and data processing. Fluorescence measurements were done in triplicate for each sample.

The fluorescence spectra of bulk samples were recorded in the emission wavelength (λ_{em}) range of 250–600 nm at excitation wavelength (λ_{ex}) in the range of 250–500 nm, spaced by 10 and 1.0 nm intervals in the excitation and emission domains, respectively. For the emission spectra of diluted samples, the λ_{em} range was from 250 to 500 nm, and the λ_{ex} range was from 250 to 400 nm, spaced by 10 and 1.5 nm intervals in the excitation and emission domains, respectively.

2.4. Determination of individual phenolic compounds by HPLC

The HPLC analyses of wine samples were carried out using a liquid chromatograph Agilent 1200 Series (Agilent Technologies, USA) consisting of binary pump, degasser, injection valve and diode array detector (DAD). Chromatographic column Nucleodur Phenyl-hexyl (250 mm × 4 mm I.D., 5 μm) was used for separation phenolic compounds. The mobile phase consisting of 1 % acetic acid (solvent A) and methanol/water (8:2 v/v, containing 1 % acetic acid; solvent B) was pumped at flow rate 1.0 mL/min. The gradient program of mobile phase was as follows: linear gradient from 0 % B to 40 % B in 12 min, held at 40 % B in

3 min, increased to 50 % B in 3 min, changed to 100 % B in 2 min, held at 100 % B in 5 min, and finally held at 0 % B in 4 min to equilibrate the column. The column temperature was maintained at 25 °C and injection volume was 20 μL. The spectrophotometric detector was set at 280 nm and spectra were recorded in wavelength interval 190–400 nm. The retention times of gallic acid, protocatechic acid, caftaric acid, (+)-catechin, caffeic acid, *p*-coumaric acid, were 7.3 min, 11.0 min, 13.5 min, 14.5 min, 16.2 min, 20.8 min, respectively (RSD < 0.9 %). The LOQ values ranged from 0.5 to 1.3 mg/L and working concentration range was from 1.5 to 50 mg/L (R^2 0.9896–0.9972) for caffeic, *p*-coumaric and protocatechic acids and it was from 1.5 to 100 mg/L (R^2 0.9921–0.9951) for gallic acid, caftaric acid and (+)-catechin.

2.5. Chemometric analysis

Data were visualized and processed with the Microsoft Excel 2018 (Microsoft Office, USA), OriginPro 2018 (OriginLab Corporation, USA) and STATISTICA version 12 (StatSoft, USA). Pre-processing of spectral data was performed using Microsoft Excel 2018. Wavelength regions containing only noise were removed from the UV–vis spectral data, resulting in 251 variables covering the ranges from 400 to 650 nm and from 230 to 480 nm for bulk and diluted samples, respectively. Pre-processing of fluorescence spectral data followed three steps of strategy described by Gonçalves Carvalho et al. [21], which deleted regions that were not related to fluorescence emission: (1) regions where λ_{em} was shorter than λ_{ex} ($\lambda_{em} \leq \lambda_{ex}$), (2) regions of the first and the second-order Rayleigh scattering, (3) regions above the second-order Rayleigh scattering ($\lambda_{em} \geq 2\lambda_{ex}$). The remaining data were unfolded into a two-dimensional matrix (number of samples × number of λ_{em} *times* number of λ_{ex}). Based on our previous work [25], a λ_{ex} ranged from 320 to 500 nm with a step of 20 nm was used for bulk samples and λ_{ex} ranged from 260 to 350 nm with a step of 10 nm for diluted wines, resulting in matrices of (49 × 1022) and (49 × 1000) for bulk and diluted samples, respectively. Then all spectral data sets were pre-processed by a mean-centering algorithm.

PLS regression [26] was done using STATISTICA version 12 (StatSoft, USA). Leave-one-out cross-validation was applied to all of the PLS regression models. Minimum root mean squared error of leave-one-out cross-validation (RMSECV) was used to select the optimum number of latent variables (LVs). The resulting PLS characteristics, number of LVs, root mean squares error of calibration (RMSEC), root mean squares error of cross-validation (RMSECV), coefficient of determination of calibration (R^2C), coefficient of determination of cross-validation (R^2CV) and relative predictive deviation of cross-validation (RPDCV) were used to evaluate and compare the predictive capabilities of the PLS models [27]. RPD values were interpreted according to Nicolaï et al. [27]: a model with RPD in the range of 1.5 to 2 can discriminate low from high values; a value in the range of 2.0 to 2.5 indicates coarse quantitative prediction; a value in the range of 2.5 to 3.0 is typical for good prediction; a value above 3.0 is reported for a model with excellent prediction accuracy.

PLS models based on UV–vis and unfolded EEM fluorescence spectral data were developed for determination of TPC in bulk and diluted botrytized wines. FC was used as reference method for the determination of TPC. The results of HPLC analyzes of wines were used in the development of PLS models for the determination of gallic, protocatechic, *p*-coumaric and caffeic acids and (+)-catechin based on unfolded EEM fluorescence spectral data of diluted wines.

In order to simplify the PLS models and possibly improve the prediction, the significant variables for the prediction using PLS were obtained by evaluating the variable importance in projection (VIP) [28] using the OriginPro 2018 software. Usually, variables with a VIP value greater than 1 are selected for the final PLS model. However, if there are many variables in the data set with VIP > 1, a higher threshold can be chosen [29]. Therefore, in addition to the VIP value of 1, other thresholds were tested in this study until no more improvement of the model was achieved. Only the best sets of variables and the corresponding PLS results are presented in this work.

3. Results and discussion

3.1. TPC by Folin-Ciocalteu method

The TPC of Slovak Tokaj selection wines, determined by the Folin-Ciocalteu method, ranged from 370 to 1223 mg GAE/L with an average of 626 mg GAE/L (Table 1). The variability among samples, expressed by the value of standard deviation (SD) of 160 mg/L (Table 1), was high compared to the SD for three replicates, which was 9 mg/L. The origin of the wines from three producers and the wide interval of years of wine production can be the reason for the high variability among samples [1,7].

The average TPC increased with increasing number of butts, for example, the relative increase for 6 butt wines was 44 % in relation to 2 butts. However, this may not be true when comparing individual wine samples, as the TPC ranges were 492–658, 370–1029, 475–895, 530–1070 and 450–1223 mg GAE/L for 2, 3, 4, 5 and 6 butt wines, respectively. In addition, the p-value slightly exceeded the significance level (α), specified as 0.05, leading to acceptance of a null hypothesis: there is no statistically significant difference between the means for 2, 3, 4, 5 and 6 butt wines. The average TPC of Tokaj selection wines was higher than the averages for Tokaj varietal wines found in the literature (mean ± SD: 490 ± 70, 526 ± 43 and 405 ± 49 mg GAE/L for Lindenblaetrige, Yellow Muscat and Furmint, respectively) [6].

Other researchers reported diverse values of TPC in Hungarian botrytized Tokaj wines. Nyitrainé et al. [30] determined 650–750 mg GAE/L in 5 butt wines and 590–670 mg GAE/L in 6 butt wines of the same vintage (1999). Compared to Slovak Tokaj wines, the average TPC was similar for 5 butt wines and significantly smaller for 6 butt wines. Pour Nikfardjam et al. [7,31] expressed the TPC as the equivalent of catechin, so the values cannot be directly compared with the results shown in Table 1. The TPC in 5 butt Hungarian Tokaj wines from vintages between 1993 and 1999 ranged from 621 to 1403 mg catechin/L (mean ± SD, 869 ± 263 mg catechin/L), lower value of 609 mg catechin/L was observed in single sample of 6 butt wine from vintage 1981. The German botrytized wines had much lower concentrations: from 248 to 747 mg catechin/L (441 ± 146 mg/l) [7].

3.2. Individual phenolic compounds by HPLC

Gallic, protocatechic, caffeic, *p*-coumaric and caftaric acids and (+)-catechin were determined in Tokaj selection wines by HPLC (Table 1). The TPC determined by FC method was significantly positively correlated with caffeic acid and *p*-coumaric acid, with R values of 0.742 and 0.693, respectively. A smaller correlation was observed between TPC and catechin, gallic, caftaric and protocatechic acids, with R values of 0.556, 0.429, −0.204 and 0.066, respectively. The contents of protocatechic, caffeic, caftaric and *p*-coumaric acids and catechin did not differ significantly for 2, 3, 4, 5 and 6 butt wines (p-value > α at α = 0.05) (Table 1). On the contrary, the gallic acid content was significantly different (higher) in the 2 butt wines compared to the remaining wines (p-value < α at α = 0.05). Catechin was the most abundant phenolic compounds in Tokaj selection wines. Other authors reported diverse values of phenolics in Hungarian botrytized Tokaj wines, mostly of an order of magnitude similar to those found in the present work (Table 1) [30–32]. Pour Nikfardjam et al. [31] reported that the most important phenolic compounds in Furmint and Lipovina (base wines for Tokaj wine production) were catechin and caftaric acid and that some of the Tokaj Aszú 5 butt wines showed high content of caftaric acid.

3.3. UV–vis and fluorescence spectra

The average UV–vis spectrum of bulk samples was characterized by

Table 1
Phenolic content in botrytized wine.

Wine	TPC (mg GA/L)	Gallic acid (mg/L)	Protokatechic acid (mg/L)	Caffeic acid (mg/L)	*p*-Coumaric acid (mg/L)	Caftaric acid (mg/L)	(+)-Catechin (mg/L)	References
Two butt	575 ± 196	**30.4 ± 8.5**	3.5 ± 2.1	11.4 ± 0.4	3.5 ± 0.5	52.6 ± 13.9	30.9 ± 21.3	This work
Three butt	599 ± 229	11.0 ± 6.3	2.4 ± 3.7	6.4 ± 2.0	2.5 ± 0.7	41.8 ± 10.6	30.4 ± 14.3	This work
Four butt	613 ± 138	11.3 ± 8.7	1.7 ± 0.5	8.3 ± 3.3	3.2 ± 0.9	30.5 ± 17.2	45.0 ± 40.3	This work
Five butt	675 ± 151	15.1 ± 7.2	2.0 ± 1.7	8.1 ± 4.2	3.3 ± 1.4	21.3 ± 33.4	41.5 ± 37.5	This work
Six butt	826 ± 259	13.3 ± 4.5	2.2 ± 2.3	8.7 ± 2.9	4.3 ± 1.3	10.3 ± 27.2	28.7 ± 16.7	This work
Mean ± SD	626 ± 160	13.1 ± 7.0	2.0 ± 1.9	8.1 ± 3.3	3.4 ± 1.3	15.3 ± 34.1	38.0 ± 31.9	This work
p-value	0.077	**0.032**	0.821	0.546	0.169	0.255	0.819	This work
Four butt	n.a.	n.a.	n.a.	n.d.	n.d.	26.4	32.9	[30]
Five butt	958 ± 333*	6.5 ± 3.4	5.1 ± 1.3	1.0	4.8 ± 1.2	12.2 ± 15.2	9.0 ± 10.0	[31]
Five butt	n.a.	10–20	1–5	0.1–1	n.a.	n.a.	1–5	[32]
Five butt	700 ± 100	n.a.	n.a.	55.8	n.d.	9.8 ± 4.4	25.2 ± 22.2	[30]
Six butt	609*	0.9	7.2	n.d.	2.6	n.d.	n.d.	[31]
Six butt	n.a.	5–10	1–5	0.1–1	n.a.	n.a.	5–10	[32]
Six butt	620 ± 100	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	[30]

* TPC given as mg catechin/L; n.a. – not analysed; n.d. – not detected; SD – standard deviation (variability among samples).

out-of-range absorbance in the wavelength range of 200 to 360 nm (Fig. 1A). At longer wavelengths, only a decrease in absorbance was observed without significant spectral features. The band at 280 nm and the shoulder in the range of 330 to 340 nm of the average UV–vis spectrum of the diluted samples (Fig. 1B) corresponded to the phenolic components (hydroxybenzoic acids as gallic and protocatechic acids, 250–300 nm; hydroxycinnamic acids as caffeic, caftaric and *p*-coumaric acids, 230–245 nm and 310–330 nm; catechins, 280 nm; flavonols, 250–270 and 350–390 nm), among others [33].

An average EEM spectrum of all bulk wines is shown in Fig. 1C. Bulk wines presented strong fluorescence in the λ_{ex} range of 390 to 500 nm and λ_{em} of 450 to 590 nm. Excitation and emission maxima were observed around 460 and 530 nm, respectively. Bulk samples had a significant absorbance at both the excitation and emission wavelengths

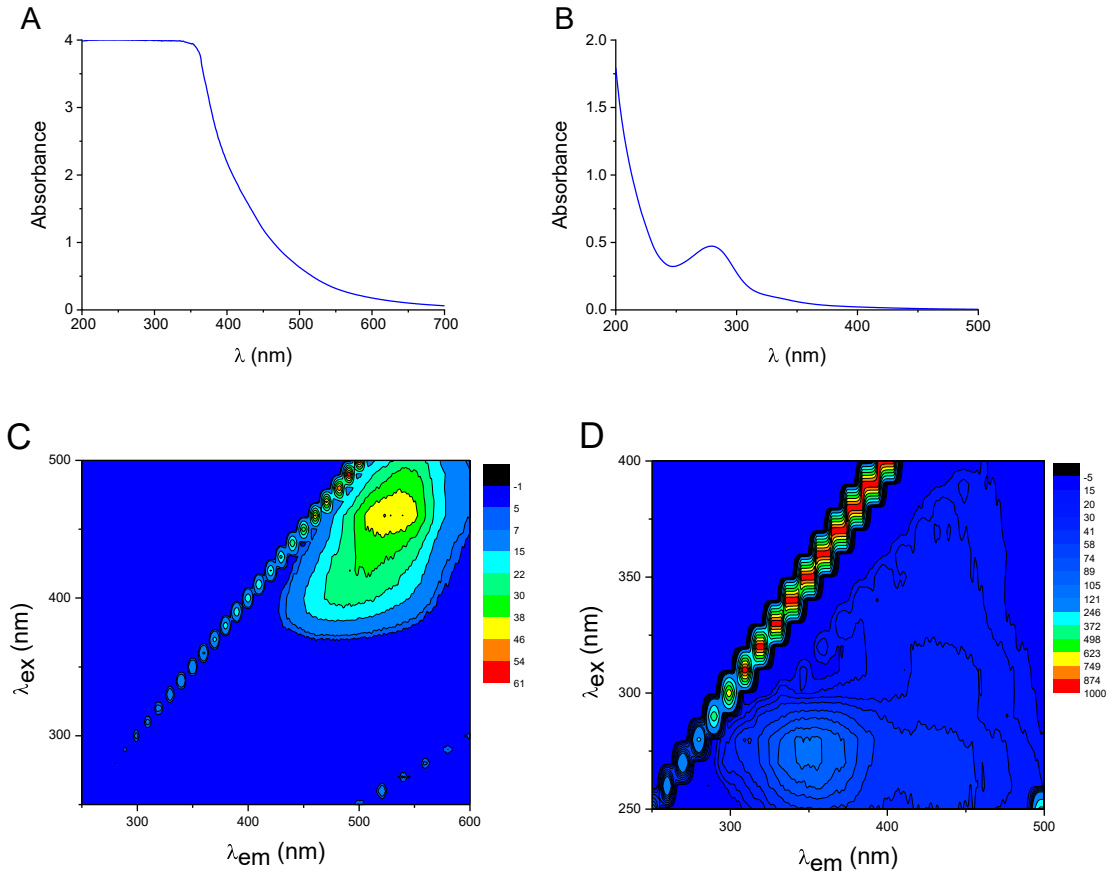

Fig. 1. Average UV–vis spectra (A, B) and contour plot of average excitation-emission matrix fluorescence spectra (C, D) of bulk (A, C) and diluted (B, D) Tokaj selection wines.

(A > 0.1 at λ < 650 nm, Fig. 1A). Therefore, the fluorescence spectra of bulk wines were affected by an inner-filter effect, which is an apparent decrease in fluorescence intensity and/or a distortion of the band shape due to attenuation of excitation beam at the point of observation and/or reabsorption of the emitted radiation [34]. According to our previous results, an inner-filter effect can be suppressed by diluting the wine samples 500 times [25]. An average EEM spectrum of all diluted wines is shown in Fig. 1D. Diluted wines presented fluorescence in the λ_{ex} range of 250 to 350 nm and λ_{em} of 320 to 450 nm. An intense band was observed at λ_{ex} and λ_{em} around 270–280 nm and 350 nm, respectively, and a weak band with excitation and emission at about 300–310 and 430–440 nm, respectively. Fluorescence spectra were similar to those previously described for phenolic acids (gallic acid, $\lambda_{ex}/\lambda_{em}$ at 280/360 nm; protocatechic acid, 270/350 nm; caffeic acid, 262,325/426 nm; caftaric acid 290, 325/440 nm, *p*-coumaric acid 290, 309/404 nm) and catechin with $\lambda_{ex}/\lambda_{em}$ at 280/310 nm [25,35,36].

3.4. TPC by PLS models of UV–vis and unfolded EEM fluorescence spectral data

PLS regression models were developed to predict TPC using individual UV–vis and unfolded EEM fluorescence spectral data recorded on bulk and diluted samples. The characteristics of the PLS models obtained in calibration and leave-one-out cross-validation are presented in Table 2.

Using all 251 variables in the UV–vis spectral ranges of 400–650 nm and 230–480 nm for bulk and diluted samples, respectively, resulted in the full PLS models with coarse quantitative accuracy as the RPD value was less than or equal to 2.5 [27]. Slightly better results, higher RPD and R2 values and lower RMSE value, were observed for the diluted samples. Fig. 2 shows VIP score plots where three wavelength regions with VIP scores >1 were identified: 400–446, 462–511 and 640–650 nm for bulk samples, and 230–244, 283–299 and 310–353 nm for diluted samples. Thus, 108 and 76 variables were selected for bulk and diluted samples, respectively, and the new PLS models were calculated on the basis of the selected variables. An increase in R2CV and RPDCV values and a decrease in RMSECV values were observed (Table 2), confirming that after selecting the most relevant variables, model complexity was reduced and prediction improved, especially for diluted samples. The RPD value of 3.8 indicated an excellent PLS model [27] for TPC determination using diluted wines. Corresponding RMSECV value of 25.2 mg GAE/L (Table 2) was lower than that described for TPC prediction based on UV–vis spectral data in red wine (RMSECV = 130 mg GAE/L) [15], white and red grape vinegar (RMSECV = 96 mg GAE/L) [17] and apple juice (60 mg GAE/L) [18]. In contrast, Longo et al. [16] reported a lower RMSECV value (20–25 mg/L) for wine juice, but using all peaks recorded at UV absorbance wavelength of 280 nm eluting between 5 and 10 min from the HPLC system to obtain TPC reference values, not the FC method. The best RPD value was 3.8 for Tokaj selection wine (Table 2). Other authors reported RPD values of 2.3, 2.7 and 3.8 for grape vinegar [17], apple juice [18] and wine juice [16].

The PLS models using all variables (1022 or 1000) of unfolded EEM fluorescence spectra provided a good predictive accuracy expressed by RPD values > 2.5, which was a better result compared to the full PLS models of UV–vis spectra. The improvement of RMSECV and R2CV values compared to UV–vis spectroscopy was also significant, mainly in the analysis of bulk samples. Because such PLS models contained a large number of variables, the variable selection method was tested, and the resulting VIP score contour maps are shown in Fig. 2. One wavelength range with high VIP scores was found for the bulk samples (Fig. 2C), with the highest VIP score (2.3) at the λ_{ex} of 480 nm and the λ_{em} of 510 nm. Using a threshold value of 1.6, (Fig. 2C), a total of 92 out of 1022 variables in the λ_{ex} range of 460 to 500 nm and the λ_{em} range of 495 to 542 nm were selected. After applying the variables in PLS, a model with R2CV of 0.956, RMSECV of 20.5 and RPDCV of 4.8 was obtained. Unlike bulk samples, the 150 significant variables for PLS modeling of diluted samples came from two wavelength regions: (1) λ_{ex} from 260 to 270 nm and λ_{em} from 330 to 420 nm with a VIP maximum at $\lambda_{ex}/\lambda_{em}$ of 270/360 nm and (2) λ_{ex} from 290 to 300 nm and λ_{em} from 360 to 420 nm with a VIP maximum at $\lambda_{ex}/\lambda_{em}$ of 300/380 nm (Fig. 2D). After selecting the most significant variables, a PLS model was obtained characterized by high values of R2C of 0.980 and R2CV of 0.972 and low values of RMSEC of 19.4 and RMSECV of 20.2 with RPDCV of 5.8.

A comparison of the results in Table 2 leads to the conclusion that the performance of the models depended on the type of spectra and samples, and the number of variables used. In all cases, VIP simplified the PLS models and improved the prediction of TPC. Considering VIP-based PLS models, RPD values were higher for unfolded EEM fluorescence spectra compared to UV–vis spectra, but the PLS model for diluted samples included more variables (150) selected from fluorescence spectra. In both spectrometric methods, diluted samples were preferred over undiluted ones. As a result, the best models had the R2CV values > 0.9, the RMSECV values ranged from 20 and 25 mg GAE /L and RPDCV values > 3.8. The number of articles on the determination of TPC in wine using fluorescence spectroscopy is limited. TPC in red wine was predicted using excitation-emission matrix in the λ_{em} range of 370 to 400 nm and λ_{ex} range of 260 to 460 nm with R2CV value of 0.77, RMSECV value of 7.4 mg GAE/L and RPDCV value of 2.1 [22].

3.5. Individual phenolic compounds by PLS models of unfolded EEM fluorescence spectral data

Using unfolded EEM fluorescence spectra recorded on diluted samples, PLS models were developed to predict gallic, protocatechic, caffeic and *p*-coumaric acids and (+)-catechin contents. The characteristics of the PLS models obtained in the calibrations and leave-one-out cross-

Table 2
PLS models based on UV–vis and EEM fluorescence spectral data for determination of total phenolic content.

					Calibration		Validation		
Sample	Spectral data	Algorithm	Variables	LV	R2C	RMSEC	R2CV	RMSECV	RPDCV
Bulk	UV–vis	PLS	251	10	0.822	60.4	0.803	63.4	2.3
		VIP-PLS	108	5	0.856	51.3	0.848	52.7	2.6
Diluted	UV–vis	PLS	251	9	0.879	45.8	0.845	47.5	2.5
		VIP-PLS	76	4	0.939	24.9	0.930	25.2	3.8
Bulk	EEM	PLS	1022	9	0.919	28.3	0.900	28.4	3.2
		VIP-PLS	92	2	0.968	20.3	0.956	20.5	4.8
Diluted	EEM	PLS	1000	6	0.865	32.0	0.861	34.0	2.7
		VIP-PLS	150	3	0.980	19.4	0.972	20.2	5.8

RMSEC, RMCESV (mg/L).

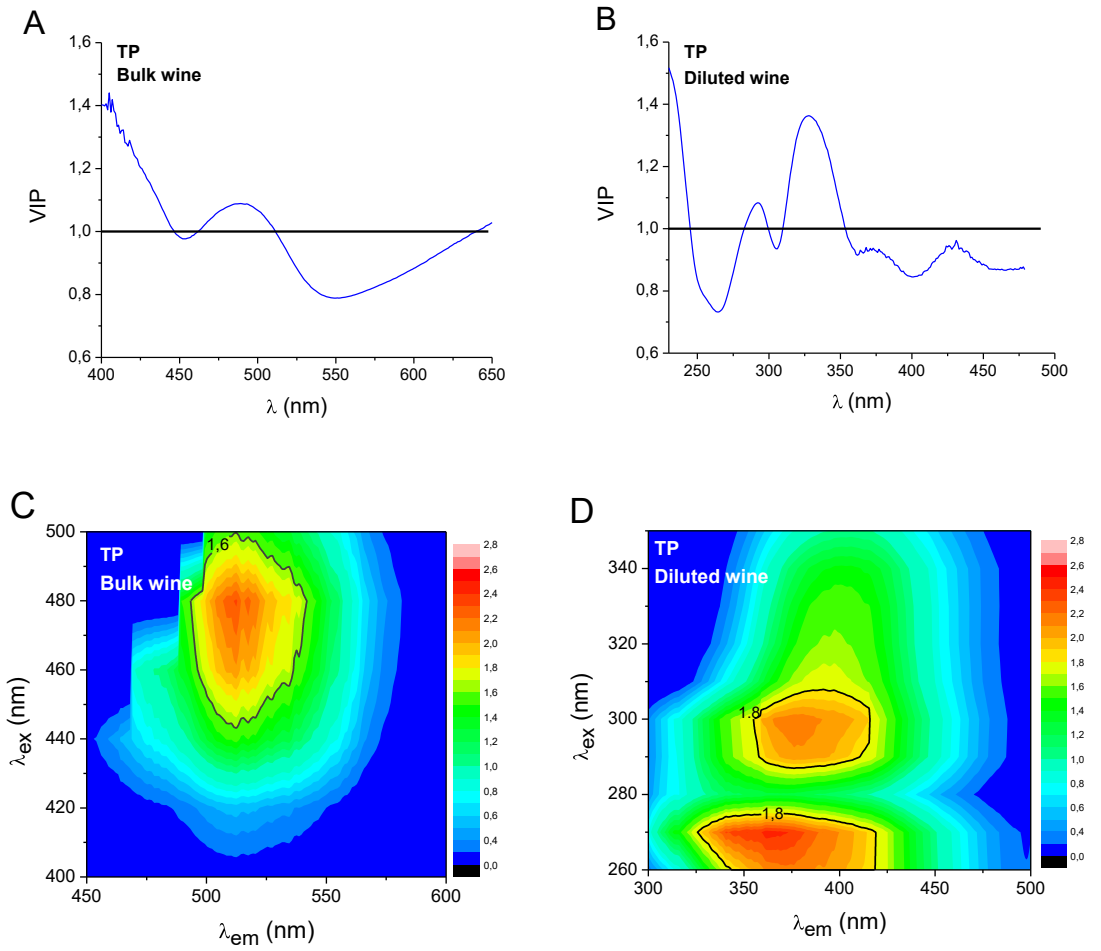

Fig. 2. Variable importance in projection (VIP) derived from the PLS model on the prediction of total phenolic content (TP) in Tokaj selection wines. (A, B – UV–vis; C, D – fluorescence; A, C – bulk wine; B, D – diluted wine). The thick black line indicates the threshold value of VIP score.

validations are shown in Table 3. Considering all five target compounds and all 1000 variables, the PLS models using the first 7–9 LVs achieved a good performance expressed by values of R2CV > 0.89 and RPD > 3.0, with the best characteristics observed for *p*-coumaric acid.

After applying variable selection method, PLS models for the four target compounds (gallic, protocatechin, caffeic acids and (+)-catechin) were calculated using 2 LVs, and one compound (*p*-coumaric acid) required 3 LVs. Compared to full PLS models, the variable selection method brought a significant decrease in the number of LVs. Additionally, an increase in R2CV and RPD values above 0.94 and 4.0, respectively, was observed. Contour plots of the VIP scores with marked threshold values are shown in Fig. 3 for the target compounds. For gallic

Table 3
PLS models based on EEM fluorescence spectral data for determination of gallic, protocatechic, caffeic and *p*-coumaric acids and (+)-catechin content.

				Calibration		Validation		
Compound	Algorithm	Variables	LV	R2C	RMSEC	R2CV	RMSECV	RPDCV
Gallic acid	PLS	1000	7	0.895	0.8	0.886	0.9	3.0
	VIP-PLS	77	2	0.957	0.6	0.961	0.7	5.0
Protocatechic acid	PLS	1000	8	0.924	0.6	0.911	0.7	3.3
	VIP-PLS	56	2	0.962	0.4	0.955	0.5	4.7
Caffeic acid	PLS	1000	8	0.951	0.3	0.940	0.3	4.0
	VIP-PLS	90	2	0.962	0.2	0.948	0.3	4.4
p-Coumaric acid	PLS	1000	9	0.964	0.2	0.956	0.3	4.8
	VIP-PLS	113	3	0.975	0.1	0.963	0.2	5.2
(+)-Catechin	PLS	1000	7	0.927	1.0	0.921	1.2	3.5
	VIP-PLS	46	2	0.949	0.9	0.940	1.0	4.1

RMSEC, RMCESV (mg/L).

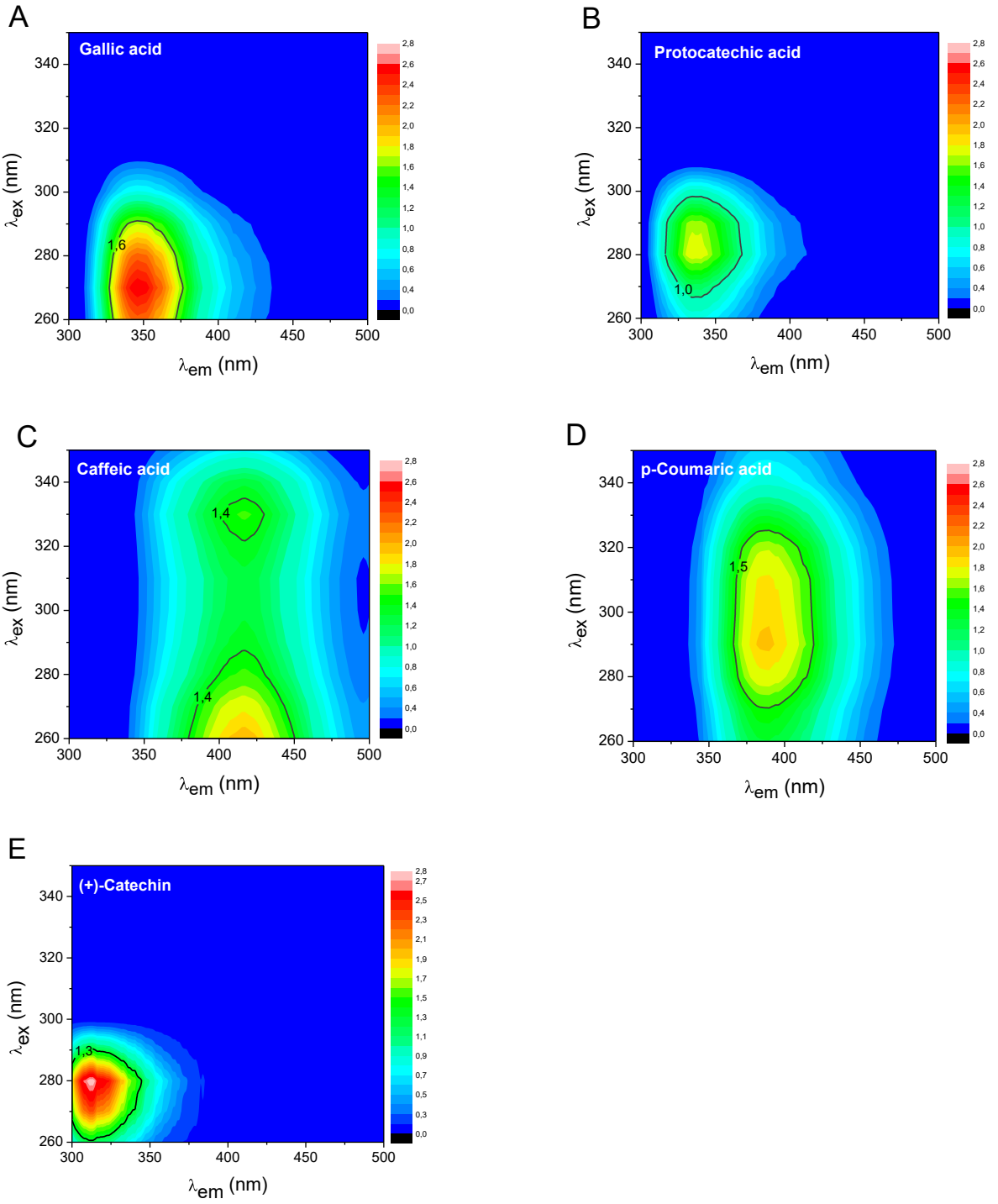

Fig. 3. Variable importance in projection (VIP) derived from the PLS model on the prediction of selected phenolic compounds. The thick black contour line indicates the threshold value of VIP score. (A – Gallic acid; B – Protocatechic acid; C – Caffeic acid; D – *p*-Coumaric acid; E – (+)-Catechin).

acid, the selected variables with a VIP > 1.6 corresponded to the λ_{ex} range of 260–290 nm and λ_{em} of 330–375 nm, and the PLS model was characterized by a high predictive ability with a RPD value of 5.0. A similar RPD value (4.7) was achieved for protocatechic acid using the selected variables with a VIP > 1.0 in the λ_{ex} range of 270–290 nm and λ_{em} of 315–365 nm. Two significant wavelength regions (VIP > 1.4) were selected for caffeic acid: (1) λ_{ex} from 260 to 280 nm and λ_{em} from 380 to 450 nm and (2) λ_{ex} at 330 nm and λ_{em} from 400 to 430 nm. This PLS model was characterized by RPD of 4.4. The PLS model with the highest predictive ability (RPD = 5.2) was obtained for *p*-coumaric acid using the selected variables (VIP > 1.5) in the λ_{ex} range of 280–320 nm and λ_{em} of 370 –415 nm. PLS model with at least good accuracy was obtained for (+)-catechin (RPD of 4.1). The selected variables were in the λ_{ex} range of 270–290 nm and λ_{em} of 300–340 nm. The cross-validation results showed RMSECV values of 0.7, 0.5, 0.3, 0.2 and 1.0 mg/L for gallic, protocatechic, caffeic, *p*-coumaric acids and

(+)-catechin, respectively. The remaining compound determined by the HPLC method, caftaric acid, was not predicted from the above fluorescence spectra as caftaric acid occurred in only about a quarter of the samples, which did not allow the development of reasonable PLS models.

The number of studies dealing with the determination of phenolic compounds in wines using fluorescence spectroscopy is limited. Previous work was focused on the use of fluorescence excitation-emission matrix method and related complex chemometric methods in the analysis of red wines. In this way, totals of anthocyanins, flavan-3-ols, hydroxycinnamates, flavonols, and totals of low molecular weight phenolic compounds (sum of all flavan-3-ols, hydroxycinnamates, and flavonols) as well as each compound separately were determined with R2 between 0.91 and 0.98 and RMSE between 0.56 and 1.4 mg/L [37], and vanillic acid, caffeic acid, epicatechin and resveratrol were predicted with R2 > 0.9 (RMSE not given) [38]. Monago-Maraña et al. [19] achieved the quantification of the content of the main phenolic compounds (catechin, epicatechin, procyanidin B, chlorogenic acid and neochlorogenic acid) in plum extracts with R2CV ranged from 0.54 to 0.89 and RMSECV from 0.05 to 0.29 mg/L using fluorescence spectra recorded at λ_{ex} of 280 and 330 nm.

4. Conclusions

The quality of a wine can be assessed using diverse physicochemical characteristics, including the TPC, as phenolic compounds significantly contribute to aroma, astringency, bitterness and color of the wine. Fluorescence spectroscopy together with chemometric analysis is a simple, fast and green alternative to the traditional FC method for the determination of TPC in wine. To the best of our knowledge, this is the first time that unfolded EEM fluorescence spectra have been combined with PLS to determine TPC and individual phenolic compounds in botrytized wines. Although the development of PLS models is time consuming in the initial stage, the subsequent determination of TPC and individual phenolic compounds is faster and cheaper compared to FC and HPLC methods. Simultaneous determination of these parameters can be important for monitoring the composition of wines during production and quality control of products.

CRediT authorship contribution statement

Michaela Jakubíková: Writing – review & editing, Conceptualization, Methodology, Formal analysis, Data curation, Visualization. **Jana Sádecká:** Writing – original draft, Supervision, Methodology, Formal analysis, Investigation, Resources, Data curation, Visualization. **Katarína Hroboňová:** Methodology, Validation, Formal analysis, Investigation, Data curation.

Declaration of Competing Interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: Michaela Jakubikova reports financial support was provided by Cultural and Educational Grant Agency of the Ministry of Education Science Research and Sport of the Slovak Republic. Michaela Jakubikova reports financial support was provided by Slovak Research and Development Agency.

Data availability

The data that has been used is confidential.

Acknowledgement

This research was supported by the Grant Agency of the Slovak Republic (VEGA no. 1/0159/20) and by the Slovak Research and Development Agency under the contract No. APVV-15-0355. Great thanks to Tokaj wineries Ostrožovič spol. s r.o., Anna Nagyová-ZLATÝ STRAPEC and Tokaj & Co., s.r.o. for the cooperation and granting the samples.

References

[1] R.S. Jackson, Specific and distinctive wine styles, in: R.S. Jackson (Ed.), Wine Science: Principles and Applications, Academic Press, Amsterdam, 2008, pp. 520–576.
[2] I. Magyar, Botrytized wines, in: R.S. Jackson (Ed.), Advances in Food and Nutrition Research, Vol. 63, Academic Press, Burlington, 2011, pp. 147–206.
[3] J. Farkaš, Technology and Biochemistry of Wine, Gordon & Breach, New York, 1988.
[4] K. Furdíková, M. Kakaš, F. Malík, Technológia výroby tokajských vín, http://www.agroporadenstvo.sk/rastlinna-vyroba-vinic-a-vino?article=755, 2015 (accessed 17 December 2015).
[5] P. Eftimov, Špecifikácia produktov pre chránené označenie pôvodu vinohradnícka oblasť Tokaj, https://ec.europa.eu/geographical-indications-register/eambrosia-public-api/api/v1/attachments/15884, 2013 (accessed 15 August 2003).
[6] D. Bajčan, A. Vollmannová, M. Šnirc, J. Árvay, J. Mezey, V. Šimanský, P. Trebichalský, R. Stanovič, M. Timoracká, Phenolic compounds and antiradical activity of Tokaj wines, J. Microbiol. Biotechnol. Food Sci. 8 (2018/19) 955–959, https://doi.org/10.15414/jmbfs.2018-19.8.3.955-959.
[7] M.S. Pour Nikfardjam, G. László, H. Dietrich, Resveratrol-derivatives and antioxidative capacity in wines made from botrytized grapes, Food Chem. 96 (2006) 74–79, https://doi.org/10.1016/j.foodchem.2005.01.058.
[8] Ľ. Ballová, Z. Eftimová, S. Kurhajec, J. Eftimová, Tokaj wines as source of polyphenols with positive effects to health, in: Investigations of Medicinal Plants – Theoretical and Practical Aspects, National University of Pharmacy, Kharkiv, 2016, pp. 4–5.
[9] M. Fikselová, M. Kačániová, M. Mellen, Antioxidant (antiradical) and antimicrobial (antifungal) efects of Slovak Tokaj wines, Acta Aliment. 39 (2010) 256–264, https://doi.org/10.1556/AAlim.39.2010.3.2.
[10] V.L. Singleton, J.A. Rossi, Colorimetry of total phenolics with phosphomolybdic-phosphotungstic acid reagents, Am. J. Enol. Vitic. 16 (1965) 144–158.
[11] A.L. Waterhouse, Determination of total phenolics, Curr. Protoc. Food Anal. Chem. 6 (2002) I1.1.1–I1.1.8, https://currentprotocols.onlinelibrary.wiley.com/doi/abs/10.1002/0471142913.fai0101s06.
[12] AOAC SMPR 2015.009 Standard Method Performance Requirements^{SM} (SMPRs) for Estimation of Total Phenolic Content Using the Folin-C Assay, https://www.aoac.org/wp-content/uploads/2020/11/SMPR202015_009.pdf, 2005 (accessed 15 June 2015).
[13] Compendium of International Methods of Wine and Must Analysis, International Organisation of Vine and Wine, https://www.oiv.int/public/medias/7907/oiv-vol1-compendium-of-international-methods-of-analysis.pdf, 2021 (accessed January 2021).
[14] C.W. Beaver, J.F. Haebertson, Comparison of multivariate regression methods for the analysis of phenolics in wine made from two Vitis vinifera cultivars, Am. J. Enol. Vitic. 67 (2016) 56–64, https://doi.org/10.5344/ajev.2015.15063.
[15] K. Skogerson, M. Downey, M. Mazza, R. Boulton, Rapid determination of phenolic components in red wines from UV-Visible spectra and the method of partial least squares, Am. J. Enol. Vitic. 58 (2007) 318–325.
[16] R. Longo, R.G. Dambergs, H. Westmore, D.S. Nichols, F.L. Kerslake, A feasibility study on monitoring total phenolic content in sparkling wine press juice fractions using a new in-line system and predictive models, Food Control 123 (2021) 106810, https://doi.org/10.1016/j.foodcont.2019.106810.
[17] C. Cavdaroglu, B. Ozen, Prediction of vinegar processing parameters with chemometric modelling of spectroscopic data, Microchem. J. 171 (2021) 106886, https://doi.org/10.1016/j.microc.2021.106886.
[18] K. Włodarska, P. Piasecki, A. Lobo-Prieto, K. Pawlak-Lemańska, T. Górecki, E. Sikorska, Rapid screening of apple juice quality using ultraviolet, visible, and near infrared spectroscopy and chemometrics: A comparative study, Microchem. J. 164 (2021) 106051, https://doi.org/10.1016/j.microc.2021.106051.
[19] O. Monago-Maraña, M. Cabrera-Bañegil, N.L. Rodas, A. Muñoz de la Peña, I. Durán-Merás, First-order discrimination of methanolic extracts from plums according to harvesting date using fluorescence spectra. Quantification of polyphenols, Microchem. J. 169 (2021) 106533.
[20] D. Gonçalves Carvalho, L. Ranzan, L. Ferreira Trierweiler, J.O. Trierweiler, Determination of the concentration of total phenolic compounds in aged cachaca using two-dimensional fluorescence and mid-infrared spectroscopy, Food Chem. 329 (2020) 127142, https://doi.org/10.1016/j.foodchem.2020.127142.
[21] D.G. Carvalho, L. Ranzan, R.A. Jacques, L.F. Trierweiler, J.O. Trierweiler, Analysis of total phenolic compounds and caffeine in teas using variable selection approach with two-dimensional fluorescence and infrared spectroscopy, Microchem. J. 169 (2021) 106570, https://doi.org/10.1016/j.microc.2021.106570.
[22] I. dos Santos, G. Bosman, J.L. Aleixandre-Tudo, W. du Toit, Direct quantification of red wine phenolics using fluorescence spectroscopy with chemometrics, Talanta 236 (2022) 122857, https://doi.org/10.1016/j.talanta.2021.122857.
[23] G. Bilge, K.S. Özdemir, Synchronous fluorescence spectroscopy combined with chemometrics for determination of total phenolic content and antioxidant activity in different tea types, J. Sci. Food Agric. 100 (2020) 3741–3747, https://doi.org/10.1002/jsfa.10413.

[24] A.d.P. Rodrigues, J.S. de Gois, M.A.J.L. Costa, C.S. da Silva, V.L. Xavier, A.S. Luna, Exploring multivariate linear regression methods for the prediction of total phenolic content in standard American lager beers using synchronous fluorescence spectroscopy fused data, Chemom. Intell. Lab. Syst. 206 (2020) 104168.
[25] J. Sádecká, M. Jakubíková, P. Májek, Fluorescence spectroscopy for discrimination of botrytized wines, Food Control 88 (2018) 75–84, https://doi.org/10.1016/j.foodcont.2017.12.033.
[26] S. Wold, M. Sjöström, L. Eriksson, PLS-regression: a basic tool of chemometrics, Chemom. Intell. Lab. Syst. 58 (2001) 109–130, https://doi.org/10.1016/S0169-7439(01)00155-1.
[27] B.M. Nicolaï, K. Beullens, E. Bobelyn, A. Peirs, W. Saeys, K.I. Theron, J. Lammertyn, Nondestructive measurement of fruit and vegetable quality by means of NIR spectroscopy: A review, Postharvest Biol. Technol. 146 (2007) 99–118, https://doi.org/10.1016/j.postharvbio.2007.06.024.
[28] C.M. Andersen, R. Bro, Variable selection in regression - a tutorial, J. Chemometrics 24 (2010) 728–737, https://doi.org/10.1002/cem.1360.
[29] X.C. Song, W. Canellas, E. Asensio, C. Nerin, Predicting the antioxidant capacity and total phenolic content of bearberry leaves by data fusion of UV-Vis spectroscopy and UHPLC/Q-TOF-MS, Talanta 213 (2020) 120831, https://doi.org/10.1016/j. talanta.2020.120831.
[30] S.D. Nyitrainé, B. Nagy, M. Kállay, Polyphenol content examination of Tokaji Aszú wines, https://ives-openscience.eu/5298/, 2014.
[31] M.S. Pour Nikfardjam, G. László, H. Dietrich, Polyphenols and antioxidative capacity in Hungarian Tokaj wines, Mitt. Klosterneuburg Rebe Wein Obstb. Fruchteverwert. 53 (2003) 159–165, https://doi.org/10.1016/j.foodchem.2005.06.014.
[32] F. Drawert, G. Leupold, V. Lessing, Z. Kerènyi, Gaschromatographische bestimmung der inhaltsstoffe von gärungsgetränken. VI. Quantitative gaschromatographische bestimmung von neutralstoffen (kohlenhydraten) und phenolischen verbindungen in Tokajer weinen, Z. Lebensm. Unters. Forsch. 162 (1976) 407–414.
[33] J. Sádecká, M. Jakubíková, Classification of Tokaj wines by ultraviolet–visible spectroscopy, Food Anal. Methods 15 (2022) 56–66, https://doi.org/10.1007/s12161-021-02097-y.
[34] J.R. Lakowicz (Ed.), Instrumentation for Fluorescence Spectroscopy. Principles of Fluorescence Spectroscopy, Springer US, Boston, MA, 2006.
[35] D. Airado-Rodríguez, I. Durán-Merás, T. Galeano-Díaz, J.P. Wold, Front-face fluorescence spectroscopy: A new tool for control in the wine industry, J. Food Compos. Anal. 24 (2011) 257–264, https://doi.org/10.1016/j.jfca.2010.10.005.
[36] C. Martin, J.-L. Bruneel, F. Castet, A. Fritsch, P.-L. Teissedre, M. Jourdes, F. Guillaume, Spectroscopic and theoretical investigations of phenolic acids in white wines, Food Chem. 221 (2017) 568–575, https://doi.org/10.1016/j.foodchem.2016.11.137.
[37] R.K.R. Ranaweera, A.M. Gilmore, D.L. Capone, S.E.P. Bastian, D.W. Jeffery, Spectrofluorometric analysis combined with machine learning for geographical and varietal authentication, and prediction of phenolic compound concentrations in red wine, Food Chem. 361 (2021) 130149, https://doi.org/10.1016/j.foodchem.2021.130149.
[38] M. Cabrera-Bañegil, M.D.C. Hurtado-Sánchez, T. Galeano-Díaz, I. Durán-Merás, Front-face fluorescence spectroscopy combined with second-order multivariate algorithms for the quantification of polyphenols in red wine samples, Food Chem. 220 (2017) 168–176.